Arm Business Briefing Q&A

March 17, 2025

Presenter: Ian Thornton, VP Investor Relations, Arm Holdings plc

Q&A

Speaker 1

Q1:

I have two questions. First, regarding the outlook for license revenue. While I understand that there can be significant quarter-to-quarter volatility, I'd like to ask how we should think about license revenue over a longer-term horizon—say one to two years, or even three years. In particular, I believe it's very important to understand how new architectures, such as Armv9, contribute to future royalty revenue. Rather than looking at license revenue on a quarterly basis, could you share your view on how we should assess and interpret license revenue from a medium-term perspective?

Second, regarding royalty revenue. Based on the current guidance, royalty revenue is expected to grow in the mid-20% range year-over-year. In the cloud space, CPU share appears to be on the rise. However, when looking at the overall spending breakdown in AI data centers, it seems that GPUs still account for a larger portion, and the monetary weight of CPUs may be relatively smaller. Given that, how much can CPU-based royalty revenue continue to drive growth going forward? In addition, I understand that automotive-grade centralized computers featuring NVIDIA's Blackwell chips are expected to launch in the second half of this year. With that in mind, I'd appreciate your thoughts on the potential upside in royalty revenue from the automotive sector from next year onward. I'd be grateful if you could address both license and royalty revenues in your response.

A1:

(Thornton) As we are still in Q4, FYE25, only a couple of weeks away from the end of the Q4, and with our results coming up in May, I will have to defer guidance for the coming year to what we provide for the information in May. But I am happy to answer your question in more general terms, in terms of the drivers of our licensing and our drivers of our royalty revenues.

You make a very good point about our license revenues being quite variable quarter to quarter. The revenue recognition, particularly around some of the large subscription deals that we sign, mean that we can end up recognizing 50, 60% of the revenue associated with a multi-year deal all on signature, and that can mean that the license revenues can be lumpy. To help with that, we do provide an annualized contract value number, and that I would always recommend that you look at first, because that basically takes the revenue recognition and spreads it out as if the revenue recognition was ratable.

If you can imagine a license that is a five-year, $20 million a year license, the revenue will

be recognized $50 million on signature, and then the remaining $50 million will be spread over the following 20 quarters, so $2.5 million per quarter. So, big spike, and then very little for the remainder of the period. Whereas with Annualized Contract Value (ACV), we will spread that out 20, 20, 20, 20, 20. So, in any quarter where you see the license revenue has gone up strongly, please go and take a look at ACV. You will probably find it is only up a little bit, and if in any quarter where the revenue is down strongly, I would suggest you look at ACV, you will probably find it is up a little bit. I thoroughly recommend you take a look at ACV.

In terms of what the drivers of our license revenue have been in recent times, and what this will mean going forward, in the past year, and particularly when I think back to the IPO, we have seen much stronger licensing than we had anticipated. In fact, if you go back to the IPO and compare to where we are guiding the market, we were originally guiding about $1.3 billion, looks like we will now be ending up closer to $1.8 billion. So, much higher.

Almost all of that is due to excitement around AI. If you are starting a new chip design today, regardless of whether that chip design is going into a smartphone, a data center, a smart TV, a car, even a washing machine, you know that that chip is going to have to run some form of AI algorithm. These tend to be quite large, quite computationally heavy algorithms, but they are also changing very rapidly.

It takes two to three years to build a chip, but the algorithms are changing quite fundamentally every six to 12 months. In that two-to-three-year period, then the model is likely to change once or twice. Anticipating what the performance requirements are going to be of that chip is very hard.

What we have seen in the past year is that companies have been buying our most advanced technology, and buying higher performance technology than we had anticipated, so that they can future-proof themselves against whatever the future design is going to be. So, really, they are over-specifying what we had anticipated, so that they can run that algorithm when their chip comes to market in two to three years' time. That has been driving our license revenues in the near term.

In terms of, if I look forward, do I expect that to continue? I cannot see it slowing down at the moment. I cannot see the demand side slowing down, because all the time, we see demand for higher performance compute to run the next models. I think, at least for a while, we will see the demand continue. How that translates into revenue will depend on when the lumps occur. So, as I say, I would still recommend that you look at ACV every quarter.

On the royalty side, one potential downside of holding a 50% market share is that we are increasingly exposed to the shifting dynamics of the semiconductor industry. And there are a few parts of the market that can have an inventory correction that will also impact Arm. We do well in the parts of the industry that are booming, but if there are any parts of the industry that is slowing down, then we get impacted by that as well. There is nowhere to hide anymore.

In the past year, we saw a slowdown in the sale of chips going into networking equipment, particularly in things like wireless communications, as there was a pause in the rollout of some 5G networks. Those were all Arm-based chips. So, there was an impact there. We have also seen an

inventory correction in industrial IoT as well. During the pandemic, many companies struggled to buy chips. And so, their supply chains were very tight. Post pandemic, they moved from being just-in-time to being just-in-case. They allowed their inventory levels to become much higher. But now, many companies believe that they have more resilient supply chains, more diversity in their supply chains. They are moving back to a just-in-time approach to inventory management, and therefore, are winding down the amount of inventory that they have, and therefore, buying fewer chips.

Again, it is not a demand issue. It is just an inventory management issue, which we hope will come to an end next year. Under normal circumstances, we would expect strong royalty revenue growth in the coming year, given that two key market segments were weak this year and we anticipate a recovery. However, that outlook is now clouded by geopolitical uncertainties and the potential impact of increased tariffs. Could that mean that there could be some form of recession? Possibly. And if that results in fewer cars and smartphones being sold, then that will be something we will not be able to avoid. That said, I think there is a lot of long-term growth drivers within the semiconductor industry that we will benefit from. Analysts project that the market will grow at just under 10% annually over the next five years. We're gaining market share and benefiting from higher royalty rates. Of course, if a recession or even a depression occurs, that growth may be more modest. But we'll have to see how things unfold as we head into next year.

Speaker 2

Q1:

Two questions. How important is Stargate Project to Arm? I would have thought you would get the CPU designs anyway in the network centers or the cloud centers. That is the first question.

Secondly, could you speak to the design and competitiveness of your CPUs, GPUs, and NPUs? What feedback are you hearing most consistently from customers in terms of areas for improvement across these three core product lines? And consequently, where are you focusing your investments to enhance future performance?

A1:

(Thornton) First of all, we are not contributing to the funding of Stargate Project. We are just going to be a beneficiary of the chips being deployed within Stargate Project. There are many decisions that have not yet been taken in terms of exactly what is going to go into Stargate Project, so I have no visibility at the moment as to what the exact mix of technologies is going to be included within there. But Arm is the CPU of choice, and so we do expect there to be a lot of Arm CPUs in there.

NVIDIA is the technology partner of choice, so my default assumption is that we'll see a significant number of Grace Blackwell chips—particularly with the Grace component being Arm-based. That said, it's difficult to predict the exact number of Grace Blackwell units or the precise product mix at this stage. Still, an initial $100 billion investment suggests a substantial volume of chips, and hopefully, a large portion of those will be Arm-based. But for now, the exact figures remain unclear.

OpenAI is going to be responsible for the management of the data center, and so it will be

down to them as to what the mix of technologies are used, and we will have to wait for them to make that determination. So not yet sure, but it certainly sounds very exciting. Sounds like it should be good, but exactly how many dollars that results into of royalty revenues, I am not sure yet.

In terms of our technology in the data center, where we are gaining share within the data center predominantly is with the cloud service providers developing their own CPU chips for use in their own data centers. Companies like Amazon developing their own Graviton chip, Microsoft developing their own Cobalt chip, and Google developing their own Axion chip. The benefit for them in developing their own chip, rather than buying an off-the-shelf chip, is that they can optimize the software that needs to be run, the workloads that they are using or that their customers are using in the data center, and they can optimize it with everything else in the data center.

For many years now, the big cloud companies have been building their own data centers and customizing the equipment that goes into those data centers, but by designing their own chips, they can also design their own blades and their own server systems. So now they can optimize the software to run on their chips, to run in their racks, to run in their servers, to run in their data centers. So, the whole thing is optimized top to bottom, left to right.

All of our major customers in this space are saying that by doing this, they can achieve somewhere between a 40% to 60% reduction in power by designing their own Arm-based chips, rather than using a traditional chip from Intel or AMD. And this is not any magic due to the Arm processor. Our low power design certainly contributes. But, the real benefit comes from optimizing the chip for a specific set of workloads. By tailoring the architecture to the application, you get a much more efficient solution. We believe this workload-specific optimization is a key driver behind the market share gains we've seen so far. Amazon announced in November last year at their AWS re:Invent conference that for the past two years, more than 50% of the new chips that they have deployed within AWS have been their own Arm-based Graviton design. Microsoft and Google's datacenter chip both hit general availability in October last year, so they are just beginning to ramp now. We see no reason why in five years' time, they could not also be at 50% or higher penetration.

The reason for that is that when Amazon first started to deploy Graviton, not all of the software that was needed to run in the cloud was available to run on Arm-based chips. Amazon initially offered Graviton to software developers free of charge to encourage adoption. That is no longer necessary, as the software ecosystem has now been fully ported. Today, there is no reason why software cannot run on Arm. And also, Amazon was targeting Graviton at their third-party customers, their AWS customers, so they could only move as fast as their customers wanted to move. And again, their customers had to port and test their software.

Now, Amazon made the decision to charge their customers 40% less running on an Arm-based chip than they did for an x86-based chip, and that is because Amazon were getting the benefits of the lower power consumption, so they were able to pass that on to their customers, so that was very encouraging. Our understanding is that Microsoft and Google are also planning to run on some internal workloads, as well as using it for Azure and GCP (Google Cloud Platform). So, the next time you have a conference call using Microsoft Teams, it might be an Arm CPU that is running the Teams for you. The next time you go to YouTube, it may well be running on an Arm-based chip,

because those are the applications that Google and Microsoft are targeting first.

More broadly, if I look across all of the major hyperscalers, of the top 10 largest hyperscalers in the world, eight of them are now deploying Arm-based chips, and the other two have their first Arm-based chips still in development, and maybe towards the end of this calendar year or early next calendar year, we would expect that to be 10 out of 10.

In terms of what they are asking from us, the key purpose of a CPU is to run software, and so the conversations that we are mainly having around next CPU designs are very much focused on what is the software that it needs to run. We are very fortunate with the relationships that we have with many of the companies that are writing the base algorithms for AI, companies like OpenAI, companies like Meta, and we are working very closely with them in terms of what is the evolution of the software that they are creating, and the dream is that you have a new CPU going into a new chip that becomes available just at the same time as the new software algorithms become available.

Now, that has not been happening in recent years because the models are changing so fast, but the more we can do a job in talking to these companies that are developing these algorithms, then hopefully we will be able to do a better job to making sure that the hardware technology is coming to market just at the time the software is coming to market, and that is definitely one of the key benefits, I think, in things like Stargate Project because it is enabling us to get more of an internal view as to what OpenAI's plans are than otherwise we would have if they were just another third-party software company.

Q2:

Just to follow up, is there any requirements from your customers you think needs to improve on the GPU, NPU side of things in particular that you think you are behind versus what the customers and the market needs in the datacenter?

A2:

(Thornton) At the moment we only have CPUs, so while I'll touch on NPUs in a moment, we currently don't have a data center product in that area, making that part of the question somewhat moot. Our primary focus is on CPUs. One consistent feedback we receive is that customers would like us to bring our CPUs to market faster. That's certainly something we're always striving to improve. However, it's a process that requires coordination with software partners, so we're limited by their development timelines as well.

As for NPUs, we do offer neural network accelerators, but these are targeted at embedded devices rather than data centers. They're designed for use cases such as robotics, security cameras, and other edge applications. In fact, one of the applications that we have been working with a company on is, how to put not large language models but a small language model even into a simple device like a washing machine.

Now you may think, why do I want a small language model in a washing machine? Well, the intention is to make the washing machine a subject matter expert in itself, so basically the manual for the washing machine gets embedded within the device itself. Then you can, rather than

just putting your clothes in and fiddling with a dial, you can actually tell it what the clothes are, what you need doing, and it will have enough of an understanding to be able to program itself for that particular wash.

Now it will not be able to tell you what the weather is going to be, it will not write you a poem, it is just going to be a subject matter in itself, but it is certainly looking at how we can take some of these AI technologies and create new product categories, things that did not exist before, and that is something that we are trying to enable with our embedded NPU.

Speaker 3

Q1:

When we look at all the research and development that Arm has done since SoftBank Group first invested, one area that stands out is not having a chip designed for AI accelerators in the data center. I was wondering if you could talk a little bit about that, why you did not do it, whether there is a possibility you do it going forward?

A1:

(Thornton) On the chip part, I mean, Arm is an IP provider, so you would not expect us to be selling chips. But if I can re-ask the question in maybe a slightly different way, which is why have we not developed an AI accelerator as an IP deliverable to license to NVIDIA and others. It is not a technology problem, it is a market demand problem. If you look across the companies such as Google and what they are doing with their TPU, if you look at what Amazon are doing with Trainium and Inferentia, what NVIDIA is doing, everyone is doing something slightly different and they are optimizing, they have strong views about how they want their algorithms to work and to run, and they are each creating something that is highly differentiated from the others.

Arm as an IP company, we do best when we are licensing the same design to everybody, or at least to multiple companies, and that way we can benefit from designing something once and licensing that same design to three, four, five companies. The goal is always, or as often as possible anyway, to try to cover the cost of the development through the license fees, and then that means that the royalties, when they turn up, are profit. But the problem is if everyone is trying to build something different, then we do not have the same thing that we can license to everybody.

What I might anticipate though is that, and certainly we have seen elsewhere in the history of the semiconductor industry, is that right now everyone is focusing on frontier models. They are trying to push the boundary of what can be done, and they are doing this by throwing lots of technology at the wall to see what makes progress, what does not, and that is leaving a lot of space behind for efficiencies. We have seen this a little bit with DeepSeek, as an algorithm which focused less on being the frontier model, but more on making an efficient version of an existing model. That was able to demonstrate, if you believe the claims, a 10x improvement by optimizing what was one of those frontier models. Historically, what tends to happen is that, in the process of solving a problem, the solution itself begins to emerge. Once the pace of algorithmic innovation slows down, it opens the door for more optimized implementations—many of which we've been

able to license to other companies. Let me give an example to illustrate the point. Imagine we've solved natural language processing—meaning, we've figured out how to enable a computer to understand language and generate natural responses. This capability could be applied in different ways: for instance, to understand a specific person's voice, enabling a digital assistant on a smartphone; or to understand any voice, enabling a cloud-based call center. Once the core algorithms are established, the focus of R&D investment typically shifts to the next frontier—whether it's self-driving cars, AGI, or other advanced challenges. Meanwhile, attention also turns to making those established algorithms more efficient and scalable, allowing many companies to deploy solutions like cloud-based autonomous call centers.

Now, then with the algorithms becoming more stable, you could then start to build chips and CPUs or GPUs or accelerators that are optimized for that particular algorithm, and in the same way we saw with DeepSeek making a 10x improvement by doing software optimizations, you can get similar levels of improvement by doing hardware optimizations as well. That is the time when you might see an IP company come along and saying, rather than everyone building their own custom accelerator, why not design it once and license it to everyone? After all, many companies are trying to solve the same problem using similar approaches.

Let me give you a recent example from the AI space. Five to seven years ago, much of the focus in AI was on image classification. You might remember the popular comparison between chihuahuas and blueberry muffins—can you tell the difference between a chihuahua's face and a muffin? That kind of task represented the state of the art in AI at the time, typically running on large-scale data centers in the cloud. You cannot buy a security camera for the home today that does not have image classification, does not have object recognition, does not have face detection, does not have face recognition. A lot of the technologies that were being run on $5 billion data centers are now being run in $50 cameras.

Similarly, five years ago, there was a lot of focus on voice recognition, and now you can have text-to-speech, live translation as an algorithm running in your smartphone. So again, a $5 billion data center going into a $500 smartphone. I think that many of what we are seeing as frontier models that we are seeing today, in a few years' time, will end up being in much, much lower-cost goods, including, as I mentioned earlier, maybe even in your washing machine.

I think the answer to your question, is not so much as, well, why do you not? But it is more like, well, when does that make sense? I think it will make sense at some point. I cannot necessarily anticipate when that will be. But we are already starting to see some of the algorithms that were being developed in the data center coming into consumer electronic devices. I think that is going to be a continuous flow. Meanwhile, the models will move on to the next problem and the next problem.

Speaker 4

Q1:

I have two questions. First, regarding CSS (Compute Subsystems), I believe there may be some competitive dynamics between Arm and custom ASIC companies. It seems like both sides are aiming

to capture a greater share of the value chain. Looking ahead, should we expect these companies to remain as partners, or is Arm aiming to expand into the custom ASIC space through CSS and capture more of this market? I'd appreciate your insights on how you see this dynamic evolving.

A1:

(Thornton) They are still customers, still partners. The reason for developing CSS is that our customer base has been changing. Historically, Arm would license its technology to semiconductor companies, to companies whose primary products are computer chips. Over time, though, that has been changing. And so, we have been finding that their customers are now wanting to build their own chips, and are coming directly to Arm to license CPUs directly from us.

In some cases, those companies will still work with ASIC companies, but they want to have the relationship directly with Arm. And the reason is that as software increasingly becomes the product that you are selling, if you are selling a smartphone, you are really selling a box that runs software. The smartphone without software is just a black box. With software, you actually have something that does something. Even cars are increasingly being sold by software, either the software in the cockpit or the software in the self-driving capability. If software is becoming the product that I am selling, then it is very important for me, the OEM, to control how that software runs. And as that software runs in the chip, I therefore need to control that chip.

Now, you may choose to acquire or develop the semiconductor design capability yourself in-house, or you might decide to use an ASIC company. Even if you choose to have an ASIC company, you still will want to have that relationship directly with Arm, because it is the Arm CPU that is running your software that is so important to your product sale. Over time, we've seen a shift: our customers are increasingly the customers of our traditional customers—in other words, more OEMs than before are coming directly to Arm. These companies, however, require something more advanced than what traditional semiconductor companies typically need. So, some companies are licensing our CSS, because that is an assembly of Arms CPUs and interconnect. It is a better starting point for their chip design.

Again, they can still go and work with an ASIC company to do the rest of the chip design. Many of the companies like Broadcom and Marvell have particular expertise around high-speed interfaces and the back-end part of the manufacturing process. Those are very hard things to do. But the design of the chip, increasingly, the OEMs want to own themselves. And so, it is not that we are competing with the ASIC companies. We are not competing with Marvell. But maybe we are providing more technology to the OEMs than had been historically the case. But this is a demand pull from their customers. So, this is because of the importance of software going into an OEM's products. We are responding to their request.

Q2:

Second, I get the impression that AI workloads are increasingly being divided between the front-end and back-end. On the front-end, we see the CPU connecting to Top-of-Rack (ToR) networking, while on the back-end, GPUs are paired with high-speed Ethernet interfaces. In this back-end

domain, NVIDIA appears to have established a fairly dominant software stack. While I understand Arm is quite strong on the CPU side, I'm interested in whether there are any efforts—by Arm or others—to challenge NVIDIA's stronghold on the back-end, particularly including networking software. Or is this an area that players have essentially given up on? I'd like to understand whether there's any movement here.

A2:

(Thornton) We are indeed very strong on the CPU side. As the alternatives to NVIDIA, I think the most credible alternatives right now are what the cloud service providing companies are developing for themselves. Google's TPU, Amazon's Inferentia and Trainium chips. These are the most credible alternatives being developed. And again, these are being developed by the cloud service providing companies because they have a better understanding of what problems are trying to be solved within their data centers, compared to the general-purpose accelerator such as the NVIDIA GPU.

I recall seeing an article by Meta, and they were claiming for running the algorithms that they wanted to run, a fivefold improvement compared to an NVIDIA GPU, because whereas the GPU is a general-purpose accelerator, they had built something to run a specific algorithm. And by again, building something to run a very narrow use case, you can optimize for the chip to run that particular piece of software.

But in terms of anyone being a third-party chip provider to go and compete with NVIDIA, I think there is space for that. I think there is probably demand for that, but I do not currently see any company that is managing to achieve that. And I think to your point that NVIDIA is becoming not just a chip company, but a solutions company providing a lot of the software as well as a lot of the other systems. Many companies that have tried to compete with NVIDIA, I think have lost because of CUDA, and the software ecosystem of software developers to support CUDA. Any competitor to NVIDIA would need to be having an effective competitiveness against CUDA as well, as well as the actual GPU.

Q3:

In the AI back-end space—specifically around accelerators and networking—what is the potential for Arm to gain share in the future? Do you see any opportunity for Arm to play a role beyond just the CPU domain in this area?

A3:

(Thornton) As I say, there is space, but right now we do not have anything to offer.

Speaker 5

Q1:

I have one question regarding short-term performance. If I'm not mistaken, Armv9's share of royalty revenue has remained at around 25% for the past three quarters. I recall there was a mid-term target to raise this to around 60–70%. Could you share some context on why the ratio has remained

flat recently, and what your strategy is for increasing it going forward?

A1:

(Thornton) The proportion of chips that are Armv9 versus Armv8, Armv7, Armv6 is entirely determined by consumers going into shops and buying things. So, it is not something that we have control over. We certainly cannot influence it other than all going out and buying more smart TVs and smartphones. It is just the mathematical outcome of people going into shops and buying phones and other devices. I think what we have seen most recently is that Armv9 has more than doubled on a year-on-year basis, but has grown in line with Armv8 on a sequential basis.

To a certain extent, I think we have been slightly surprised by the growth of Armv8. And had Armv8 not grown, Armv9 would have a higher proportion. But Armv8 has been growing partly because we have seen a recovery in industrial IoT chip sales. I have mentioned earlier that there had been a decline in industrial IoT earlier in the year. We had a bit of a pop-up in Q3, and most of that is actually Armv6. It is very old technology, and so that helped to balance things out.

But also, you have to bear in mind that Armv9 is still at the moment only in high-end smartphones and in data center chips. So, everything else is still Armv8 and older. There is plenty of room for Armv9 to grow into. We are still very confident that in time, 60% to 70% of our royalty revenues will come from Armv9. And actually, in the last quarter, I think we have been just very pleased to see that Armv9 has grown, but also so has everything else. It all pays royalties. We do not really mind that much where the royalty comes from.

Speaker 6

Q1:

I have two questions. Looking at slide 18, we see that over the past two years, both revenue and headcount have grown at the same pace—17%. Going forward, do you expect this balance to shift over the next two years? If I recall correctly, the IPO plan projected that revenue growth would begin to outpace headcount growth. Could you share when you expect that inflection point to occur?

A1:

(Thornton) Actually, our revenue growth is higher than we had said at the IPO. We are exceeding our IPO targets. At the IPO, we also indicated that we thought that our non-GAAP operating profit would grow at 1 to 2 percentage points per year. We did not mention the headcount growth, but the intention was always to continue to grow investments in R&D and also to continue to grow our engineering headcount. I think we are very much on track.

I think looking forwards, as I mentioned, I am not going to be giving revenue forecasts today, but we certainly intend to keep investing by hiring new engineers. I think right now there is a lot of opportunity as we see more algorithms coming from the data center to the edge. We need to make sure that runs on Arm technology.

Also, we have a partnership now with OpenAI through Stargate and through Cristal

intelligence. We absolutely must take this opportunity to make sure that we are developing the right technology, CPU or accelerator, so that we can maximize the opportunity for us in the data center as well as outside of the data center. We have every intention to continue to grow our headcount at the same sort of trajectory. We added a thousand engineers last year. We will add another thousand engineers next year at a minimum.

Furthermore, we may start doing some of that hiring through acquisitions. To date, we have done all of our hiring organically, but historically we have had a history of buying small semiconductor companies. We do not want their technology, so we are not buying them for their products, but we are buying them for their engineering teams.

These can be 200, 300 people companies. It is great because that is one whole CPU team. It comes with graduates, senior managers, project managers, program managers. Often it comes with a building. We can just go in there, take away the stationery, give them all Arm business cards, and away they go as an Arm design team. That is much faster than hiring people one at a time. That is something I think you can expect to see, but right now we have no plans to slow down hiring. This is not the time to be focusing on profitability. I think even if revenue ended up being a lower number than we had originally anticipated, I think that we will continue to hire through that, because I think the long-term opportunity is so great.

Q2:

I believe there hasn't been any change to your dividend policy, but could you please provide an update on your current thinking around dividends and your future dividend policy?

A2:

(Thornton) To date, we do not pay a dividend, but we do have plenty of cash. We have about $2.5 billion of cash and cash equivalents, and we are probably going to be increasing that by approximately $1 billion per year. We have only recently done an IPO, so doing a big buyback is probably not a good idea. Also, we get many complaints about the lack of liquidity or the lack of float, so that is not a desired outcome. Other than some small acquisitions, but they will not absorb our cash, and so therefore that leaves the dividend.

So today we do not pay a dividend. I think SoftBank Group, as our much larger shareholder, would have to want to desire a dividend, but we are very interested in getting feedback from other investors, and also SoftBank Group's investors as well to see whether a dividend is required. We are very much open to a dividend if there was sufficient demand for one.

Speaker 7

Q1:

I have two questions. First, as a follow-up to an earlier question, I'd like to ask about which business layers Arm intends to focus on going forward. Recently, there have been comments suggesting that Arm is no longer just operating behind the scenes in chip design, but is also moving into adjacent layers—sometimes as a partner, and at times as a competitor. Looking ahead, how do you see this

partner-versus-competitor dynamic evolving? Will Arm continue to strengthen its role as a partner, or as the company adds more value, do you expect to increasingly compete with players in those adjacent layers? Could you share your long-term view on customer relationships and how Arm's position across different business layers may shift over time?

A1:

(Thornton) I believe the relationship will continue to be one of strong partnership. However, as I mentioned previously, our customer base is expanding, and we are seeing more companies who traditionally bought chips wanting to take control of the chips that go into their products, and they want Arm to provide the technology directly to them. They want for Arm to provide more technology than we have licensed to semiconductor companies.

Now, maybe some of our customers would rather that we did not provide technology directly to them, but I think it is up to them to demonstrate that they can add more value, that they can build a chip for Amazon better than Amazon can build their own chip. And if they cannot, then it is only appropriate that Amazon have the choice of developing the technology themselves.

So, I would not say that we are competing with ASIC companies. Rather, we are enabling their customers to become more independent if they so choose to do so. But I think we are providing more choice rather than we are competing with our customers.

Q2:

My second question is about Cristal intelligence. Could you please explain the relationship between Arm and Cristal intelligence? Since Cristal intelligence is expected to be developed by OpenAI, I would assume it will be based on Arm architecture. In that sense, will Arm technology be integrated into Cristal intelligence itself? Additionally, there have been comments that SoftBank Corp. will use Cristal intelligence internally to support management decisions etc. once it's available. From Arm's perspective, does your involvement with Cristal intelligence extend to actually using the technology once it's launched, or is the relationship focused purely on the design side? I'd appreciate hearing your expectations and outlook regarding Arm's role in the Cristal intelligence project.

A2:

(Thornton) I think there is two very exciting parts of this. At one level, I see Cristal intelligence being a little bit like Stargate Project. It is an opportunity for Arm-based chips to be deployed within a large high-value data center. So hopefully, lots of chips, lots of opportunity for Arm-based chips. But in addition, as part of this, we will get access to the OpenAI agents, and we will be able to use those internally to improve our own business processes, and also potentially our own product development processes as well.

As I mentioned earlier, Rene (Rene Haas, CEO, Arm Holdings plc) has indicated that the SG&A headcount will remain effectively flat. As the engineering organization continues to grow, we'll be expected to support a much larger team—without a corresponding increase in staffing. To meet this challenge, we need to prioritize automation, including AI and agentic AI, to scale our

capabilities and effectively support a significantly larger engineering organization over time. On the engineering side, I think that we have experiments going on to see whether how AI could be useful in our product development. We have been using Microsoft's GitHub Copilot as part of our software engineering activity for some time. The reports I have heard is that it has had a roughly 10% productivity improvement, although it is quite patchy between different people. Some people get more, some people get less. But that is over 2,000 people. So, another 200 people's worth of work we have been able to achieve through that.

As it is quite patchy, I suspect that therefore, as we get used to using the technology more, then there is a lot more room for a lot of productivity gains. To date, we have not rolled out the AI on the CPU design side of things, but I am sure that is something that we will be looking to do so in the future. And hopefully, that would also have a similar benefit in improving CPU design going forward.

Speaker 8

Q1:

Around the longer term, if we are on the journey to AGI and ASI, that suggests exponential pickup. What do you see your kind of segment mix looking like when we get to those two groundbreaking moments?

A1:

(Thornton) Well, if we get to ASI, who knows? Maybe we will all be communicating by telepathy and all have chips embedded in our brains. I have no idea. I always find it very interesting to hear Masa speak about the opportunity for AGI or ASI in terms of having all vehicles being autonomous, and there never being an accident again. I mean, the amount of computes that would require in each vehicle will be much greater than anything we have obviously seen so far. You need to have multiple cameras, LiDAR, radar, vehicle-to-vehicle communications, vehicle-to-infrastructure communications, as well as obviously to the cloud communications, as well as having a smart brain in the center.

We are working at the moment with many of the car companies to make that happen. We are working with about 120 of the world's leading car companies as part of a consortium focused on defining the software requirements and APIs necessary to enable the software-defined vehicle. That is work ongoing, but still, it is probably many years away before that becomes a reality.

But then you also start thinking about what that level of automation could do in factories, in the home, in restaurants for preparing food. I mean, all of that will be powered by maybe a super brain that sits in the cloud, but to get things done physically, you are going to have a lot of local sensors, cameras, controllers, actuators to move the robot's limbs around, and its fingers and joints and things like that. I think all of these markets are going to become much larger if we get to a world of AGI or ASI. Arm already now has a very high share of the controller chips, and the brain, and the camera chips that go into all of these devices. So, I think this looks very positive for us if we do end up in a world like that.

Speaker 9

Q1:

I'll keep it brief, as many questions have already been asked. First, Oracle has been mentioned as either an investor or a business partner in the Stargate project. Could you please explain the nature of Arm's relationship with Oracle in the context of hyperscaler-type AI data center business? For example, you mentioned earlier that Arm has a 50% share with AWS—what is the situation with Oracle? As inference becomes an increasingly important part of AI workloads, how do you see the potential for Arm to increase its wallet share with Oracle? Also, the initial Stargate investment appears to involve Oracle's data center in Texas—could you share any insights or updates on Arm's involvement or performance in that deployment?

A1:

(Thornton) Unfortunately, regarding Stargate Project and the investments, we are not involved at all with the investing in Stargate. So, I have no answer for you. Maybe the SoftBank Group IR team might be a better place to answer that question. Regarding our relationship with Oracle, again, they are an investor in Project Stargate. So, they are helping to fund many of the Arm-based chips. So, thank you, Oracle. Thank you, Larry. But for us, I think the main relationship is through Ampere.

Ampere is a semiconductor company that Oracle and Arm are both investing in. But significantly, Ampere is providing a lot of the Arm-based chips that Oracle is deploying. So, the relationship we have is with Ampere, but Ampere's customer is Oracle. That is how that relationship is established.

Q2:

This may be a question you've been asked in various settings already, but as inference costs decline—particularly with technologies like DeepSeek—do you see this as a positive development for Arm's demand outlook? In other words, does improved cost efficiency increase elasticity, potentially leading to stronger demand growth? Or do you see it having the opposite effect? I'd appreciate hearing your thoughts on this.

A2:

(Thornton) As I mentioned earlier, I think DeepSeek, to a certain extent, should not have surprised anyone. When you focus on the frontier models, very much focused on performance and trying to do new things. But if you start to focus a little bit more on efficiency, and take those frontier models and then try to create a more efficient version, you can make a large amount of progress very quickly, which I think is probably what has happened with DeepSeek.

I think there should have been less surprise about DeepSeek. I think some of the surprise was more to do with the fact that it was a Chinese company. Had it been a U.S. startup, I think everyone would have been less startled by it. To a certain extent, I think there was a lot more efficiency opportunities to come. I mean, because that was just an efficiency focusing on software. Once the algorithms stop moving around, there is the opportunity to focus on the hardware side of

things as well. And you can probably see another 10x improvement when you start developing the hardware to accelerate the software.

As I indicated earlier, we can see this happening in how some of the image classification models that were looking at pictures five years ago, the image classifications of chihuahuas and muffins are now in security cameras going into the home. When you stop focusing on the frontier and start focusing on optimization, then you can make substantial efficiency improvements, which obviously therefore enables edge devices. And so, yes, we absolutely expect to see a large language model functionality going into edge devices. And at the edge, it will be running on Arm processors. So yes, all very good for Arm.